# MP2

For every file stored in our system, it has 4 replicas. When client receives command "Put", it will send a request to master to get four ips of datanode that file can be put on. After master send certain ips to client, client will send a request to that ip to upload the file asked. In the end, when upload finishes, datanode will send an ack to client. Master will mark this task as done until all datanodes send acks to client. When client receives command "Get", it will send a request to master to get an ip of datanode that contains file we want. After master send certain ip to client, client will send a request to that ip to download the file asked. In the end, when download finishes, client will send a ack to master to tell master the task is done.

Test get
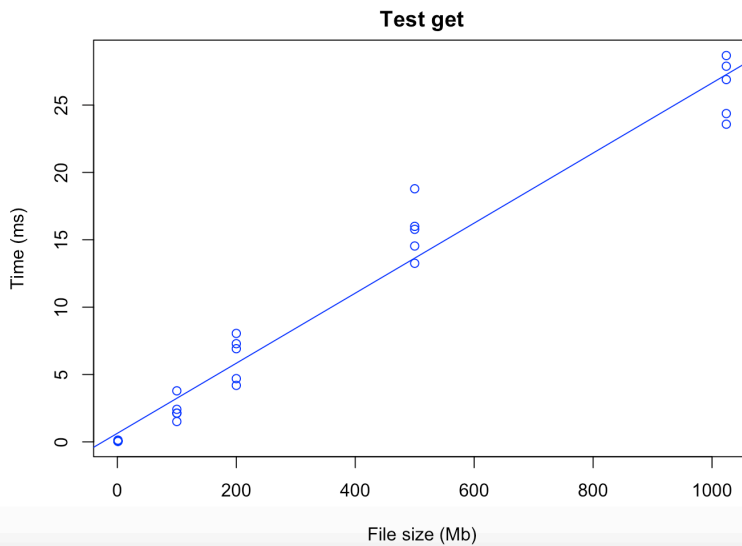
get	1M	100M	200M	500M	1G
mean	0.07	2.39	6.22	15.27	26.27
sd	0.04	0.84	1.68	2.05	2.21

Test Put
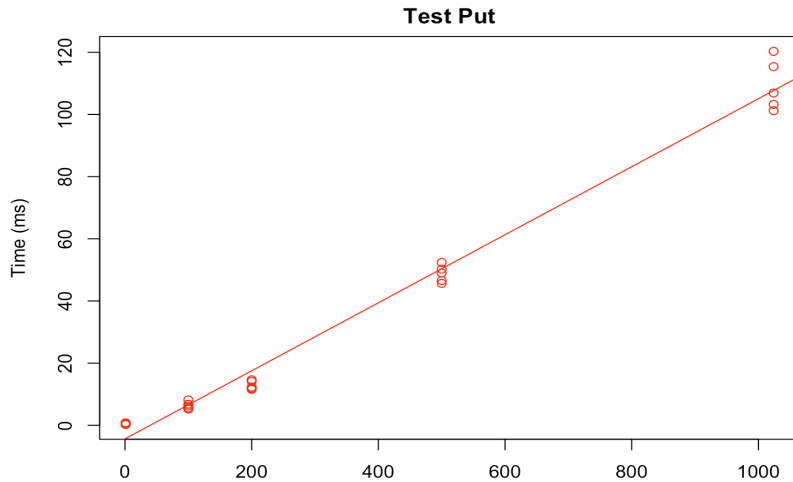
put	1M	100M	200M	500M	1G
mean	0.48	6.39	12.89	48.76	109.39
sd	0.2	1.15	1.4	2.71	8.15

Test Wikipedia
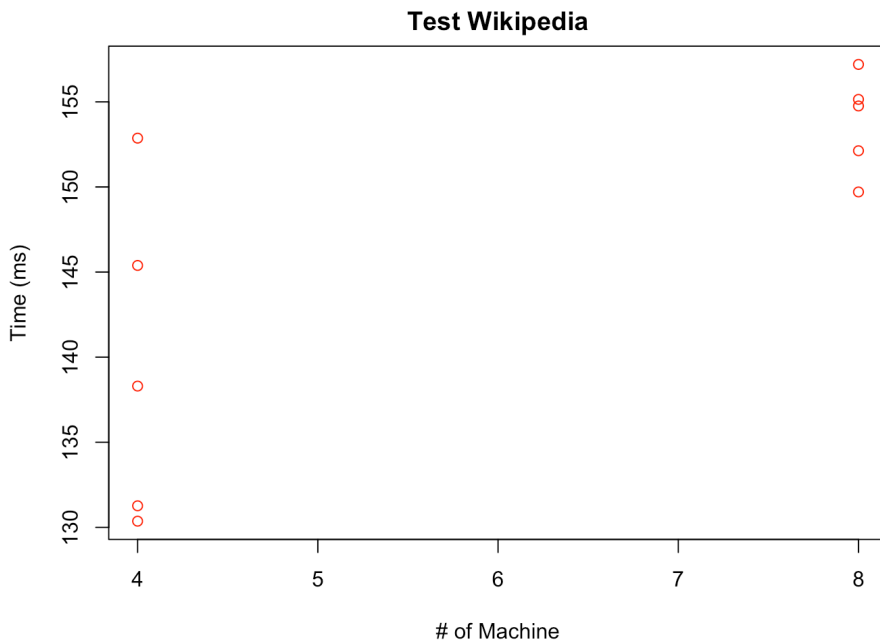
# of Machine

wikipedia	4 machines	8 machines
mean	139.63	153.78
sd	9.56	2.9

Our system performs just like we expect Because larger file size means longer io time leads to longer time for this system to complete storing or getting. Put needs about 4 times longer than Get, since, for each file put, our system needs to upload 4 replicas.